The imperative for regulatory oversight of large language models (or generative AI) in healthcare

Year: 2023 | *Citations: 668* | *Authors: B. Meskó, E. Topol*

Abstract

The rapid advancements in artificial intelligence (AI) have led to the development of sophisticated large language models (LLMs) such as GPT-4 and Bard. The potential implementation of LLMs in healthcare settings has already garnered considerable attention because of their diverse applications that include facilitating clinical documentation, obtaining insurance pre-authorization, summarizing research papers, or working as a chatbot to answer questions for patients about their specific data and concerns. While offering transformative potential, LLMs warrant a very cautious approach since these models are trained differently from AI-based medical technologies that are regulated already, especially within the critical context of caring for patients. The newest version, GPT-4, that was released in March, 2023, brings the potentials of this technology to support multiple medical tasks; and risks from mishandling results it provides to varying reliability to a new level. Besides being an advanced LLM, it will be able to read texts on images and analyze the context of those images. The regulation of GPT-4 and generative AI in medicine and healthcare without damaging their exciting and transformative potential is a timely and critical challenge to ensure safety, maintain ethical standards, and protect patient privacy. We argue that regulatory oversight should assure medical professionals and patients can use LLMs without causing harm or compromising their data or privacy. This paper summarizes our practical recommendations for what we can expect from regulators to bring this vision to reality.